



Open Science Grid

Document Name	Open Science Grid Deployment, Spring 2005
Version	0.9.2
Date last updated	Feb 14 th 2004

Abstract. We present the plan for the initial deployment of the Open Science Grid (OSG) in the Spring of 2005. We refer to this set of functionality as OSG Release 0.2. As a persistent, evolving grid, the OSG will assimilate and extend the current Grid3 in its initial deployment. We expect subsequent deployment activities to build off this base deployment, continue to support backward compatibility and add services driven by the needs of the stakeholders.

1	Open Science Grid Environment.....	1
1.1	OSG Release 0.2 Elements:	2
2	Organizational Infrastructure.....	3
2.1	The Deployment Activity.....	4
2.2	The Integration Activity	4
2.3	Service Deployment Process.....	5
3	OSG Deployment Plan	6
3.1	“Flipping the Switch” between Grid3 and OSG.....	7
4	OSG Release 0.2 Services	7
4.1	Compute Elements (CE).....	7
4.1.1	Resource Allocation Policy	7
4.2	Storage Elements (SE).....	7
4.3	OSG Operations.....	8
4.3.1	Documentation and Publication	8
4.3.2	Incident Response	9
4.4	OSG Discovery Service.....	9
4.5	Organization Membership Services (VOMS).....	9
4.6	OSG Software Release	9
5	References.....	10
	Appendix 1: Lists of Organizations and Facilities/Sites	10
	Appendix 2: Service Readiness Plan	10

1 Open Science Grid Environment

The initial deployment of Open Science Grid strives to build upon the Grid3 success and relax several of the restrictions imposed on that collaboration. In particular, the Open Science Grid will contain services of various release levels, resources which service only a subset of all OSG VOs, and an increased membership pool.

Open Science Grid will include the continuing sites of Grid3, together with additional existing DOE and University facilities and campus Grids. The initial deployment of Open Science Grid will support an application mix from the US LHC collaborations, LIGO, SDSS, BTeV, CDF, D0, BaBar and STAR experiments, GADU, SnB from ACDC, FMRI at Dartmouth, other bio-informatics applications such as BLAST, as well as computer science demonstrator applications. Open Science Grid governance documents describe the process by which Organizations join. That process is being ramped up slowly at first to debug it and avoid swamping the system. This document describes the technical requirements and milestones for OSG deployment.

The Open Science Grid environment will be an evolution of the Grid3 common grid infrastructure. The main goals of the evolution are:

- To continue and improve operation of a reliable production Grid for users while incorporating new resources.
- To take the next steps towards a widely used production Grid for our science Stakeholders.
- To further demonstrate interoperability between the OSG and LCG infrastructures, and demonstrate initial interoperation with the TeraGrid.

Open Science Grid is a “marketplace” for resource providers and consumers to barter for exchanges of value. As such, it defines a minimal set of common services and interfaces and allows for participants to form any combination of partnerships consistent with the charter. Members of the OSG consortium may provide software to implement services which conform with these requirements. Effort is taken to collect useful, consistent sets of this software into OSG maintained sets, however there is no requirement that participants run common software – rather they must provide conformant service interfaces. Thus an OSG “Release” refers to a set of functionality and agreed interfaces rather than the collection of software which provides exemplars of this functionality and interfaces. This statement of principle is tempered with the practical observation that many service definitions and interfaces are not yet stable enough for bulletproof specification, so members are strongly encouraged to use common software where possible. One of the desirable features of the “marketplace” model is that the consumers will determine what level of conformity they demand and which features are most desirable.

The initial deployment is planned for Spring 2005 and will include a number of additional or extended services to those on Grid3. The final list will depend on integration testing, but we have identified the following as Required (we will hold off deployment until these are ready), Highly Desirable (priority efforts are being made to get these ready, but they will not stop deployment), Desirable (identified as goals for the Spring Deployment)

Required:

- OSG Governance documents defining:
 - What is required to “be in” OSG.
 - What is the process for joining or leaving OSG.

- What are the allowed uses of the OSG.
- Grid3 CE resources which choose to continue to run as is.
- Configuration process for OSG release resources.
- CE resources running recommended VDT version including:
 - gatekeeper with authorization callout
 - authorization stub to authorization service
 - identity mapping stub
- Definition of required monitoring system.
- Storage Elements with SRM v1.1 interface.
- Operations infrastructure sufficient to support the above at least at the level of the Grid3 experience:
 - OSG GridCat instance
 - OSG MonALisa infrastructure
 - OSG Ganglia collector
 - Basic Support statements for how to obtain support

Highly Desirable:

- Functional Security Incident Response and Handling system.
- OSG-wide Discovery Service.
- New Applications
 - US LHC analysis applications
 - Tevatron Run II, Star and BaBar simulation jobs
- Managed access to OSG resources from LCG job (elements)
- OSG-EGEE agreement on proxy authorization attribute meaning
- Compute Elements able to make authorization decisions based on proxy attributes
 - Role-aware Grid Identity to local identity mapping services
 - VOMS servers and clients distributed to participating VOs
 - VDT 1.3+ generation GRAM
- SRM based CE-wide tactical storage services in "standard" CE config
- Strategic Storage elements with SRM v1.1 interface
- Common accounting mechanism to collect VO usage statistics across OSG

Desirable:

- Grid exerciser utilizing the Discovery service to inventory/test OSG

The following principals of Grid3 will be maintained:

- Experiments must be able to effectively interoperate and run their applications on non-dedicated resources.
- Applications must be able to install themselves dynamically, thereby imposing minimum requirements on grid facility managers.
- A grid architecture consisting of facilities (e.g., execution and storage sites), services (e.g., a prototype operations center, an information system for resource publication and discovery, and so on), and applications.

1.1 OSG Release 0.2 Elements:

The Open Science Grid Release 0.2 is comprised of the following common service elements:

Compute Elements (CE)

- Gatekeeper with GT2.4 GRAM compatible interface

- Validated as usable by at least 1 OSG Organization
 - Expected scale is minimum O(10) CPUs per CE
 - The specific validation is determined and documented by the Organization

Storage Elements (SE)

- A SRM v1.1 compatible interface
- Validated as usable by at least 1 OSG Organization
 - Expected scale is minimum of O(1) TB per SE
 - The specific validation is determined and documented by the Organization

OSG Operations

- Indiana iGOC is central point of contact for OSG operations
- Organization Support Centers coordinate with each other and the iGOC for OSG operations and support.

OSG Discovery Service

- Clarend Discovery Service is the common Web Service registry for OSG

Organization Membership Service (VOMS)

- Each Organization required to maintain a VOMS service
- Organizations required to have documented membership registration process

Other services will be deployed on OSG, in fact, we expect Organizations will build substantial systems atop this base, but this common set defines the minimum functionality of OSG Release 0.2. Each of these services has a detailed description of requirements later in this document.

Limitations of the Open Science Grid Release 0.2 (it is anticipated these will be addressed in future Releases):

- Data management services (include replica services) are the responsibility of the VO. In the initial deployment of OSG we are unable to support common data management services.
- There will be no grid-wide scheduling services or brokering of the workload management.
- Policies will, in general, be manually described and implemented by the VOs and/or Sites. The exception is that the priority of job execution on the sites will be through the batch system schedulers.

2 Organizational Infrastructure

The organization infrastructure for Open Science Grid includes Technical Groups and Activities, both of which contribute to the overall program of work needed for deployment. The few broadly scoped Technical Groups coordinate and oversee a technical or programmatic area. Technical Groups such as Security and Support Centers are responsible for gathering the requirements and overseeing the work done in a particular technical domain. The many Activities provide the organizational framework for the contributions and tasks of the program of work. Equivalent to a “project,” an activity will, in general, live long enough to meet a set of pre-defined deliverables, then disband. Once an Activity is defined it is expected that the stakeholders will provide the effort needed to successfully meet the deliverables, subject to the oversight of the appropriate Technical Group(s).

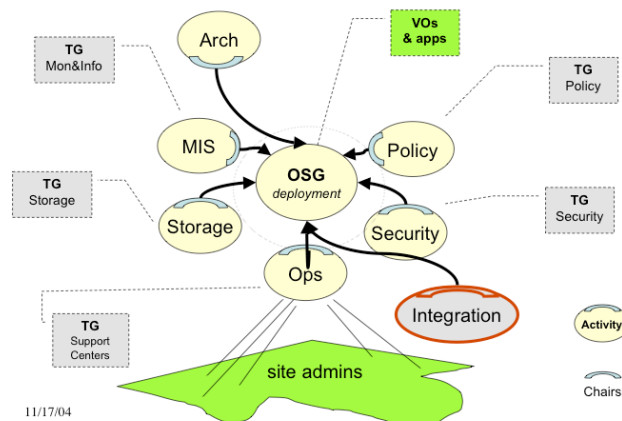
2.1 The Deployment Activity

The Deployment Activity defines and coordinates the program of work based on the requirements of the stakeholder contributors. The Deployment Activity is responsible for delivery of the first deployment of the Open Science Grid. The OSG Deployment Activity includes:

- The co-chairs of the Activities contributing to the initial deployment of the Open Science Grid.
- Liaisons from each of the VOs contributing applications to run on the Open Science Grid.
- Liaisons from each of the VOs providing Facility and Campus Grid infrastructures which will interface to and integrate with the Open Science Grid.
- Liaisons to the partner grid infrastructures – the LCG, EGEE and TeraGrid.

The scope of the deployment activity includes the deployment coordination role and is where recommendations and guidance from the blueprint group are reconciled with the deliverables and schedules. Maintaining technical coherence and making the strategic technical decisions and planning (e.g., which services to deploy when) are included in the responsibilities.

The initial relationship of Technical Groups and Activities in the Deployment Activity is shown below:



2.2 The Integration Activity

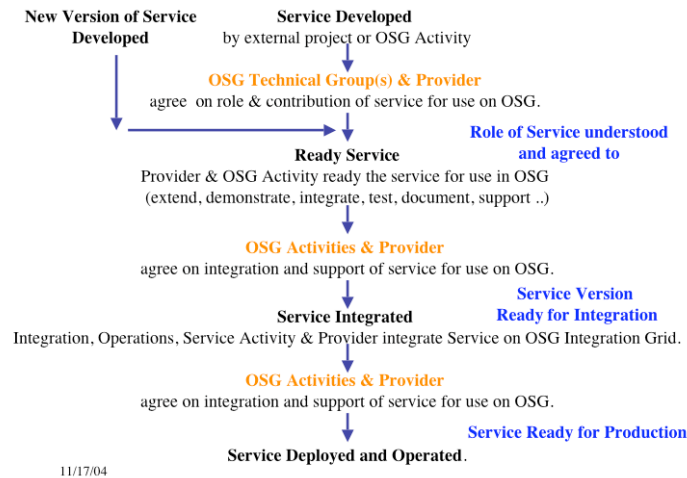
The Integration Activity is responsible for the integration and validation of the services to be deployed, and for testing consistency, performance and robustness of the common infrastructure.

The Integration Activity has a unique role. It provides the coordination of and support for validation and testing of services, technologies, and applications before they are deployed into the stable production infrastructure. Services and applications are initially tested and used in parochial testbeds or grids outside of OSG. Once they are to be deployed on OSG they must be tested in an integrated environment to ensure they are robust, operate at the scales needed, can be operated and supported in a production environment, and provide the necessary documentations and tests.

For example, if we want to add a Storage Element or Service, we expect that the service comes with a way to install it, a description of what the expected functionality is, and a method to validate the installation to confirm that the installation was a success. We expect that the service comes with a test harness that exercises all the functionality. We want to be able to say that it passed all the intended functionality at the intended scale: of the order of 10 files, 1000 files, of the order of 10 users, 1000 users,, whatever is agreed to between the developers and proponents of a service and the applications that will use it.

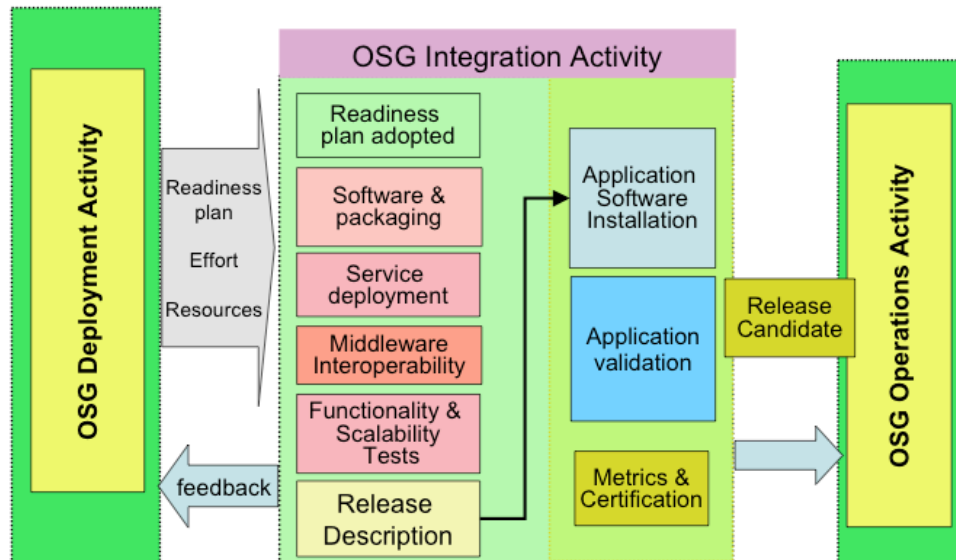
2.3 Service Deployment Process

Services and VO based applications are proposed and deployed on OSG through a lightweight process and series of steps designed to provide for robustness and maintainability of the infrastructure and prevent denial of service and performance breakdown of the grid.



A short readiness plan is presented for each service to be deployed. A draft of this plan is included in Appendix 2. The process through integration is shown below:

Service Path (WIP)



3 OSG Deployment Plan

The Open Science Grid deployment activity is anticipated to repeat on approximately yearly timescales as new collections of functionality improvements warrant a coordinated effort to evolve the OSG. As mentioned before, at any time, there will be “legacy” elements of OSG from the previous Release, those for the current generation common functionality, and specialty elements with limited scope or “warranty”. The Spring 2005 deployment of OSG, as the first of these release cycles, seeks to put in place the supporting framework for this cycle as well as the particular elements of this release.

The plan is currently:

Jan – Feb, 2005: construct an OSG operations infrastructure from a combination of new instances of services and migration of Grid3 operations activities. Collect and

Feb 2005: Complete Integration test assessments for elements proposed for Spring 2005 common functionality set. Determine whether the Required functionality can be met, or what actions need to happen. Recommend to Deployment activity the elements for the OSG release.

March 2005: Move Grid3 services to OSG operations. “Flip the switch” with OSG.

March – April, 2005: Install or migrate resources to OSG Spring 2005 “release”.

April 2005: Perform operations tests with major applications

For the Spring 2005 release, the “legacy” release is the Grid3+ release as documented <http://www.ivdgl.org/grid2003/>.

3.1 “Flipping the Switch” between Grid3 and OSG

At some point the evolution of Grid3 to OSG will be “done”. In order for this to occur the current Grid3 production grid must be reconfigured, the Site Administrators, Operations Staff and VOs must be aware and understand the change. A detailed list of steps is included in Appendix 4.

4 OSG Release 0.2 Services

This section includes the detailed discussions of requirements and configuration for the OSG Release 0.2 basic services.

4.1 Compute Elements (CE)

Two versions of Compute Elements are expected to be deployed on OSG Release 0.2: Grid3 Compute Elements and OSG Release 0.2 Compute Elements. The former are defined on the Grid3 site at <URL>, we deal only with defining the latter here.

Requirements for OSG Release 0.2 Compute Elements (CE)

- Gatekeeper with GT2.4 GRAM compatible interface
- Validation by at least 1 OSG Organization (see above)
- Monitoring interface(s) ???
- *Authorization* via Gridmapfiles or Authorization Services.
 - Sites may use GUMS for dynamic mapping of user PKI certificates to local accounts, invoked from the site gatekeeper.
 - Sites may adapt SAZ to impose site authorization and access policies.
- VOs are expected to implement policies to prevent overload of the headnodes.

Recommendation for OSG Release 0.2 CE's

- Applications should use the Grid Monitor interface and submit jobs using Condor-G, to mitigate the performance and scaling issues with the GT2 gatekeeper interfaces on the head node of each site.
- An alternative may be provided by the emerging Condor-C interface to job execution.

4.1.1 Resource Allocation Policy

For resource allocation policy will be enforced through batch queue priorities as for Grid3. Additional information relating to policy such as <http://gridline.uchicago.edu/%7Ecldumitr/data/up.status.latest.up.nicer.txt> will be made available.

4.2 Storage Elements (SE)

Storage Elements (SE)

- A SRM v1.1 compatible interface
- Validation by at least 1 OSG Organization (see above)

Persistent Storage accessible at specific sites through direct agreement between the Site and the VO.

Managed *durable and transient storage services* available at some sites accessed via SRM & GridFTP interfaces from the Grid infrastructure and Posix I/O at the application interface. More than one implementation will be supported for OSG sites - SRM/dCache, SRM/DRM(LBNL) SRM/Nest. Sites will be allowed to decide on the implementation depending on the size and

nature of the site, the VOs that have assured access to the site, and the support agreements with the technology providers. (Space management will be provided for the \$DATA and \$TMP areas of Grid3.)

Permanent Storage:

Permanent storage will be provided by Storage Elements with SRM interfaces. BNL, Fermilab and LBNL already have such installations. The work to be done includes providing the information about the services, incorporating them into the monitoring frameworks, and agreements between the storage resource providers and the VOs who will use them.

Transient Managed Storage:

The primary motivation is simply to provide application and site administrators with a rational way to reserve and allocate storage resources to the Grid3/OSG user community at the VO level of granularity. The current technologies expected to be deployed are:

1. DRM: An SRM implementation from LBNL. It can be deployed on top of a normal unix type file system (local or NFS). It is already part of the VDT as an optional install, and has been tested outside of its development environment to some degree.
2. dCache+SRM: This is a virtual file system type implementation with SRM.
3. NEST

4.3 OSG Operations

The Operational Model will be an extension of the iGOC to include a distributed set of Support Centers. A full operations plan is being developed in parallel with the Deployment Activity.

- *Information services* based on MDS and MonaLisa. The GIIS architecture will be VO specific, but a top level GIIS will be maintained to which all top level VO GIIS should report. The plug-compatible GIIS service from the LCG - the BDII – will be used in some instances to provide information filtering services, and increase the robustness of the overall information provision.
- *Monitoring* through the use of Ganglia, MonaLisa and MDS. The GridCat service will be used to publish the status of and information of all the sites on the Grid.
- Additional services and configuration tools to improve the end to end throughput and success rate of the workload (e.g. end to end traceability of all jobs, diagnostic and validation scripts). An extension to the Site_Verify script will be provided. Additional *Operational Support Services* will be provided.
- *Accounting* through a mix of the ACDC monitoring, an extension to the MDViewer application, and extended local Facility accounting services.
- The security and operational infrastructure will be extended by a requirement for signature of an Acceptable Use Policy by all users, an extension of the Grid3 site agreement at least in the area of security incident response.

4.3.1 Documentation and Publication

Information about each activity and the technologies deployed will need to be available from the Document Repository.

4.3.1.1 Security

The following documents and agreements are being developed for OSG

Incident Response Plan

Acceptable Use Policy – for users, and services (including resource providers).

4.3.2 Incident Response

Operations team will need to maintain the incident response mailing lists and support infrastructure for the adhoc incident response teams.

4.4 OSG Discovery Service

OSG Release 0.2 includes the Clarens Discovery Service as a common Discovery Service to be used by all Web Services based OSG services. This is not a hard requirement at this time, since there are few agreed web services interfaces, but it is strongly encouraged and it is expected that registry with the OSG Discovery Service will be a requirement in future OSG Releases.

4.5 Organization Membership Services (VOMS)

GRID3 VO Services are based on VOMS, with scripts on each host pulling information from the VOMS servers to the local host and a script to generate the gridmapfile. This requires each HOST to register with the VOMS service by providing their host certificate. USATLAS uses the LCG LDAP server which requires no authentication. This mechanism will continue to be supported for the OSG. However, for OSG not all VOs will be given access to each site. It will be the responsibility of the Site Administrator to set up their site correctly and ensure they only enable the VOs they support.

For OSG, site application of authorization and prioritization for processing and storage facilities will be supported for roles and groups within a Virtual Organization (VO). Simple identity and VO role based authorization and prioritization will be provided for processing and storage facilities. For processing this is through use of the Globus Gatekeeper callout. For storage is will be through the SRM callout. VOMS will continue to be used as the repository to map users to VOs.

Each site will be expected to provide software for the local call-outs. Sites implementing the extended VO services will require a version of VDT that supports GT V3.2 and some V3.3 services.

It will be important to provide good site and VO administrator documentation.

- VOs may use VOMRS for registering and managing the user database.

4.6 OSG Software Release

- *basic grid middleware* based on the Condor and Globus Toolkit GT2 services within the Virtual Data Toolkit (VDT) version 1.3 or greater. This includes GT3.2. The recommended version of Globus GRAM is the pre-web services version.
- Support for Pacman (2 and 3) for the *packaging, installation and configuration* of the common grid infrastructure. The core recommended software packages will be available through central pacman caches, together with example configuration scripts.

5 References

1. Open Science Grid web site <http://www.opensciencegrid.org>
2. Grid2003 web site <http://www.ivdgl.org/grid2003>
3. Samgrid web site <http://projects.fnal.gov/samgrid>
4. Grid2003 Plan http://www.ivdgl.org/grid3/documents/document_server/uploaded_documents/doc--707--Grid3_v21.pdf
5. Grid2003 Lessons Learned
http://www.ivdgl.org/grid3/documents/document_server/uploaded_documents/doc--751--Grid2003%20Project%20Lessons-6.doc
6. Teragrid <http://www.teragrid.org>
7. LHC Computing Grid <http://lcg.web.cern.ch/LCG/>
8. EGEE <http://www.cern.ch/egee/>
9. Grid User Management System (GUMS) <http://grid.racf.bnl.gov/GUMS/>
10. Site Authorization Service (SAZ) <http://www.fnal.gov/docs/products/saz/SAZ.htm>

Appendix 1: Lists of Organizations and Facilities/Sites

VOs & Applications

The VOs contributing applications to run on the Open Science Grid deployment in Spring 05 are the represented by the following liaisons:

Liaison	Name	Applications
LIGO	Kent Blackburn	
D0	Joel Snow	
STAR	Jerome Lauret	
Babar		
CDF		
SDSS	Neha Sharma	
ATLAS	Razvan Popescu	
CMS	Greg Graham	
BTeV	Dan Engh	
MRI	James Dobson	
GRASE	Mark Green	
GADU		
GLOW	Nate Mueller	
CS-GriPhyN	Mike Wilde	

Appendix 2: Service Readiness Plan

The following is a simple outline to follow for preparing services for the OSG integration testbed.

- * overall (brief) description of the service, and expected duration of the service.
- * dependencies and interactions with other services
- * resources required -- are clients/agents/info providers etc, whatever, to be installed on all sites.

* What are the server requirements (is a central discovery server /repository needed/provided, etc)?

* how is it packaged. install -and- config instructions

* test harness -- that is, once installed, how does one understand everything is working properly. are all required functionality tests successful?

* scalability tests - what are the expectations in terms of performance and at what scale and in what metrics for successful integration.

* contact information for support of the service in the integration phase (as distinct from production operations in the OSG)

* links to documentation

A one or two pager is all that is required.

.